EyeDetect Hybrid Multi-Issue Comparison Test (HMCT)

Development and Validation Summary

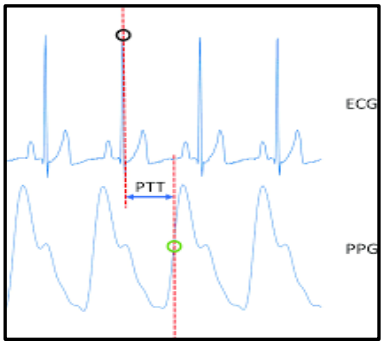
June 28, 2021

Converus recently introduced the EyeDetect Hybrid Multi-Issue Comparison Test (HMCT). The HMCT is similar to the ocular-based MCT and Audio MCT (AMCT). It covers up to four relevant issues, computes the probability that a test subject was deceptive to each of the relevant issues, and it uses probabilities to classify the test subject as truthful to all of the relevant issues or deceptive to one or more of them. The HMCT differs from all prior deception detection technologies in that it combines traditional polygraph testing methods with ocular-motor methods.

Testing with the HMCT takes place in two phases. The first phase is like a traditional polygraph test but is automated. The computer presents test questions to the test subject with a text-to-speech (digital) voice over headphones while it records electrodermal, respiratory, cardiovascular, and ocular-motor activity. The second phase consists of two sessions of the ocular-based MCT protocol. Whereas the polygraph phase of the test assumes that deception is primarily associated with emotional arousal, the MCT phase assumes that deception is cognitively more demanding than telling the truth. In the MCT phase of the test, an automated script encourages the subject to read and answer True/False statements presented on the screen as quickly and accurately as possible. This test uses time pressure to increase the cognitive demands of the task.

As with other digital polygraphs, electrodermal recordings are obtained from disposable electrodes attached to the palmar surface of fingers of the non-preferred hand, and respiration is recorded with a strain gauge around the chest or abdomen. The EyeDetect+ 2.0 station also records the electrocardiogram (ECG) from electrodes attached to each arm and peripheral vasomotor activity with a photoplethysmograph (PPG) from a finger on the non-preferred hand. From the ECG and PPG signals, the computer derives a measure known as Pulse Transit Time (PTT). PTT is the difference in time from when the heart beats to the occurrence of the pulse in the finger (see inset).

PTT has been shown to vary inversely with changes in blood pressure (Geddes et al., 1981; Obrist et al., 1978; Obrist et al., 1979). As blood pressure increases, the time it takes the pulse to travel from the heart to the periphery decreases. While changes in PTT may be on the order of only 5 ms, research has shown that PTT is at least as diagnostic as the traditional cardiograph (Webb & Kircher, 2005). EyeDetect+ 2.0 replaces the cardiograph with PTT. PTT is less invasive than the cardiograph, alleviates subject discomfort, and offers more flexibility in test construction. the duration of polygraph charts has been limited by the cardiograph since it must be deflated approximately every five minutes to restore circulation to the test subject's lower arm. Replacing the cardiograph with PTT, the test can include more test questions in each session than a traditional polygraph test. More test questions produce more measurements of physiological reactions, and the additional measurements might improve the reliability and accuracy of the test.



The HMCT introduces a new test format for the polygraph-optimzed phase of the test. Test questions are arranged in all possible pairs of the relevant issues. In the first half of the first session, questions about the first relevant issue (R1) are paired with questions about the second relevant issue (R2). In the second half of the first session, questions about the third relevant issue (R3) are paired with questions about the fourth relevant issue (R4). In subsequent sessions, R1 is paired with R3, R1 is paired with R4, R2 is paired with R3, and R2 is paired with R4. During any given sub-session, test subjects focus on only two of the four relevant issues.

The auditory presentation of Yes/No test questions every 22 seconds was optimal for recording diagnostic changes in traditional polygraph measures, such as peak amplitude of the skin conductance response, increases in blood pressure, and respiration suppression. In the latter sessions, the cognitive demands of reading and answering quickly and accurately was optimal for ocular-motor measures, such as number of fixations, increases in pupil diameter, and blink rates. Features extracted from the polygraph-optimized sessions and ocular-motor optimized sessions were combined by means of a binary logistic regression equation to classify test subjects as either credible or not credible on the test and on each of the relevant topics individually.

On cross-validation, HMCT decisions were 90% correct when it classified the test subject as truthful to all relevant questions or deceptive to any one or more of the relevant questions. HMCT decisions were 91% correct on individual relevant topics.

The remainder of the present paper describes the procedures used to develop the Hybrid MCT and estimate its accuracy.

Experiment

We conducted a mock crime experiment modeled after Cook et al. (2012) to develop and cross-validate a statistical model of polygraph and ocular-motor measures that produced a credibility score for each relevant issue. In the present experiment, there were four issues: (1) theft of cash, (2) theft of a ring, (3) theft of a cell phone, and (4) theft of a set of AirPods®.

We recruited 182 subjects from the local community. Two people were dropped from the sample due to inadequate ECG or PPG recordings. Subjects were told that some participants would commit one or more of the thefts, whereas others would be innocent and would not commit any of the crimes. Subjects were randomly assigned to one of three groups. One group of guilty subjects stole $20 from a secretary's office (n=60). They lied about taking the $20, but told the truth about the diamond ring, cell phone, and AirPods. Another group of guilty subjects stole $20 from the secretary AND stole a ring from a desk drawer (n=60). They lied about two of the four relevant issues. The third group of subjects was innocent of all four crimes (n=60); they were truthful about all four of the relevant topics covered on the test.

After subjects completed their assigned tasks, they were given the HMCT. They sat at a desk in front of a computer screen. The test proctor attached skin conductance, ECG, PPG, and respiration sensors to the subject and then calibrated the eye tracker. A Physiocom GP12 recorded skin conductance and

respiration at 340 Hz and ECG and PPG at 1000 Hz. A Tobii 4C remote eye tracker attached to the bottom of the computer monitor recorded bilateral gaze position and pupil diameter at 60 Hz.

The polygraph-optimized phase of the HMCT contained a set of 22 Yes/No questions, repeated three times. A text-to-speech, digital voice presented test instructions and questions over headphones. Subjects were instructed to avoid moving, keep their feet flat on the floor, and look at the computer screen during the test.

Table 1 shows the test questions for the first of three polygraph sessions. Note that the first sub-session asked about the R1 and R2 issues, whereas the second sub-session asked about the R3 and R4 issues. Sessions two and three paired issues R1 with R3, R2 with R4, R1 with R4, and R2 with R3. Prior to each sub-session, subjects were informed about the two relevant issues that would be covered in that sub-session, e.g., *Now you will be asked about the diamond ring and the twenty dollars.*

Table 1. Test questions for the first polygraph session

Question Number	Sub-session	Question Type	Question
1	1	Neutral	Is looking both ways before crossing the street a wise thing to do?
2	1	R2	Are you guilty of the theft of the diamond ring?
3	1	R1	Are you guilty of the theft of the twenty dollars?
4	1	R2	Did you steal the diamond ring?
5	1	Neutral	Does it take about 180 days for the earth to revolve around the sun once?
6	1	R1	Did you steal the twenty dollars?
7	1	R2	Are you innocent of stealing the diamond ring?
8	1	R1	Are you innocent of stealing the twenty dollars?
9	1	Neutral	Do large trucks get better gas mileage than newer compact cars?
10	1	R2	Did you take the diamond ring?
11	1	R1	Did you take the twenty dollars from the wallet?
12	2	Neutral	Have you read at least one book, magazine or newspaper article in your life?
13	2	R3	Are you the one who stole the cell phone?
14	2	R4	Did you take the AirPods from the Toyota Camry?
15	2	R3	Are you guilty of stealing the missing cell phone?
16	2	Neutral	Does Facebook let you stay in touch with friends and family?

17	2	R4	Did you steal the AirPods from the red Camry?
18	2	R3	Did you steal the cell phone from the office?
19	2	R4	Are you innocent of the theft of the AirPods?
20	2	Neutral	Do McDonald's, Wendy's and Burger King serve hamburgers and fries?
21	2	R4	Are you the one who stole the AirPods?
22	2	R3	Are you innocent of the theft of the cell phone?

The ocular-motor optimized phase of the HMCT followed the three polygraph sessions. The computer informed subjects that they should read and answer the statements quickly and accurately or they might fail the test. Thirty-two True/False statements were presented twice in different orders; eight statements addressed each of the four relevant topics. In addition to the physiological and oculo-motor measures, the computer recorded response times and the number of questions answered incorrectly. Table 2 contains the test statements in the first repetition of the ocular-motor portion of the HMCT.

Table 2. Test questions for the first ocular-motor session

Statement Number	Statement Type	Statement
1	R3	I admit to stealing the missing cell phone.
2	R1	I did not steal the secretary's twenty dollars.
3	R2	The diamond ring is missing from the desk drawer because of me.
4	R4	The AirPods went missing because I took them.
5	R3	I did not steal the cell phone from the office.
6	R1	I am the person who stole the twenty dollars.
7	R4	I did not steal the AirPods from a car in the parking lot.
8	R2	The diamond ring in the desk was stolen by me.
9	R4	The missing AirPods were stolen by me.
10	R1	The reason the $20 is gone is because I stole it.
11	R3	I am the person who stole the cell phone.
12	R2	I stole the diamond ring from the desk drawer.
13	R3	I am innocent of the theft of the cell phone.

14	R2	I did not take the diamond ring from the desk.
15	R4	I am innocent of the theft of the AirPods.
16	R1	I admit to taking the stolen twenty dollars.
17	R4	I am guilty of the theft of the AirPods.
18	R2	I did not steal the diamond ring from the desk
19	R3	The cell phone was missing because I stole it.
20	R1	I am not the person who stole the $20.
21	R4	The AirPods were not stolen from the car by me.
22	R1	I am innocent of stealing the twenty dollars.
23	R3	I did not steal that missing cell phone.
24	R2	I am innocent of the theft of the diamond ring.
25	R3	The missing cell phone was not stolen by me.
26	R2	I admit to stealing the diamond ring from the desk.
27	R4	I stole the AirPods from the red Camry.
28	R1	The cash in the wallet was not stolen by me.
29	R4	I did not take the AirPods from the Toyota Camry.
30	R1	The secretary's twenty dollars were stolen by me.
31	R3	The missing cell phone was stolen by me.
32	R2	The diamond ring in the desk drawer was not taken by me.

Subjects were paid $40 for their participation and were offered an additional $30 bonus if they passed the test.

Analysis

Ocular-motor data were analyzed to identify features that discriminated between questions and statements answered truthfully and deceptively. A subset of polygraph and ocular-motor features was identified that achieved over 90% accuracy on the complete set of relevant questions. For each subject, those features were weighed and combined by means of a logistic regression equation that generated a credibility score for each relevant question. The credibility index was the probability that the person was

truthful about that topic. If the credibility index was 0.5 or greater, the subject was classified as truthful to questions about that issue. If the credibility index was less than 0.5, the subject was classified as deceptive about that issue.

K-Fold Validation

A statistical model that is optimal for classifying the cases in a particular experiment is rarely optimal for the population from which the subjects were sampled. The model is not optimal because the sample does not perfectly represent the more general population from which it was drawn. Consequently, the model produces biased estimates of accuracy if it is tested on the cases used to create the model.

Better estimates of accuracy can be obtained with k-fold validation. A k-fold validation divides the data set into *k* folds (subsets). The first subset comprises a hold-out subsample and is removed from the dataset. The remaining subsets are combined to create a training set. A logistic regression model is developed using the cases in the training set. That regression model is then used to classify the cases in the hold-out subsample. The accuracy observed in the hold-out sample provides a less biased estimate of accuracy because the holdout cases were not used to optimize feature coefficients in the regression equation. The accuracy achieved in the hold-out sample is recorded.

This process continues for each partition of the data set. The first subset is returned to the training set, and the second subset is removed to serve as a new holdout sample. A new logistic regression equation is created with all but the second subset of cases. That model is used to classify cases in the holdout sample, and its accuracy is recorded. This process is repeated for each of the remaining subsets. The best estimate of accuracy for the model is mean accuracy across the *k* holdout samples.

Validation of the HMCT

Each of 180 subjects was truthful or deceptive to each of four relevant questions. This provided a total of 720 (180 X 4) relevant questions where the person was truthful or deceptive. The sample of 720 questions was split into 6 subsamples (folds) of 120 questions. Most of the relevant questions were answered truthfully. Innocent subjects answered all four relevant questions truthfully. Guilty Cash subjects answered three of four relevant questions truthfully, and Guilty Cash+Ring subjects answered two of four relevant questions truthfully. The high percentage of relevant questions answered truthfully (75%) was equally represented in each of six folds. Ninety of the 120 questions in each subsample were questions answered truthfully, and 30 were questions answered deceptively.

For each fold, logistic regression provided a new combination of polygraph and ocular-motor measures that was optimal for relevant questions in the training set. That regression equation was used to classify cases in the holdout sample. The cutoff probability was set at the base rate of deception ($p = .25$). Table 3 reports the accuracy (percent correct) for questions answered truthfully or deceptively for each fold. Accuracy estimates ranged from 73% to 100% correct. Mean accuracy was 90.8% across all folds.

Table 3. Percent correct decisions for questions answered truthfully or deceptively in 6-fold validation

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Mean
n	120	120	120	120	120	120	
Truthful	92.2	87.8	94.4	93.3	93.3	92.2	92.2
Deceptive	93.3	96.7	93.3	80.0	86.7	91.7	90.0
						Mean Accuracy	**91.1**

On average, accuracy was slightly higher for questions answered truthfully (92.2%) than for questions answered deceptively (90.0%). At the level of individual relevant questions, mean accuracy on cross-validation was 91.1%. Based on these results, *we would expect the HMCT to produce 91% correct decisions on individual relevant issues when the model is used with a new sample*.

The posterior probabilities of truthfulness for subjects in the holdout samples were used to classify subjects as either truthful to all four of the relevant questions or deceptive to any one or more of the relevant questions. The subject was classified as innocent if the posterior probability of truthfulness was .50 or greater on all four relevant questions. Otherwise, the subject was classified as guilty. Of the 60 innocent subjects, 54 were correctly classified (90.0%). Of the 120 guilty subjects, 110 were classified correctly (91.7%). On cross-validation, mean accuracy was 90.8%. *Based on these results, we expect the HCMT to classify 90.8% of subjects correctly in a new sample of cases.*

References

Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S., Woltz, D. J., & Kircher, J. C. (2012). Lyin' Eyes: Ocular-motor Measures of Reading Reveal Deception. Journal of *Experimental Psychology: Applied, 18(3)*, 301-313.

Geddes L.A., Voelz, M.H., Babbs C.F., Bourland J.D., & Tacker W.A. (1981). Pulse transit time as an indicator of arterial blood pressure. *Psychophysiology, 18,* 71–74.

Obrist, P. A., Light, K. C., McCubbin, J. A., Hutcheson, J. S., & Hoffer, J. L. (1978). Pulse transit time: Relationship to blood pressure. *Behavior Research Methods & Instrumentation, 10,* 623-626.

Obrist, P. A., Light, K. C., McCubbin, J. A., Hutcheson, J. S., & Hoffer, J. L. (1979). Pulse transit time: Relationship to blood pressure and myocardial performance. *Psychophysiology, 16,* 292-301.

Webb, A. K. & Kircher, J. C. (2005). *Use of Pulse Transit Time for the Psychophysiological Detection of Deception.* Final report to the U.S. Department of Defense. Salt Lake City: University of Utah, Department of Educational Psychology.